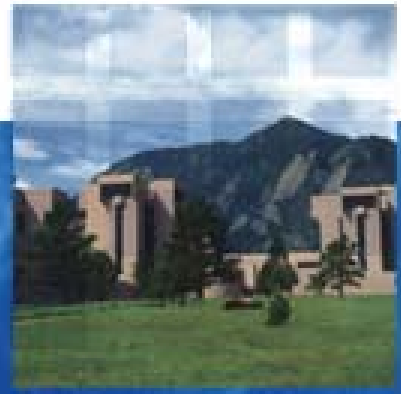
NCAR



Aura Data Guidelines
A Common Approach For All Instruments

Cheryl Craig

National Center for Atmospheric Research

Why discuss Aura Approach?

- Aura experience demonstrates that it is possible for developers at diverse locations creating multiple products from several instruments to agree on important common data product formats
- Common formats substantially simplify end user's tasks when combining data observed by one or more instruments
- Common formats simplify the task of developing both the data sets themselves and of tools for manipulating them.
- Provides lessons learned from a successful effort to develop common standards over multiple instruments at distributed locations
- Provides insight as to time and effort to produce a set of data guidelines

What are the Aura Guidelines?

- A set of Guidelines explicitly defining format of similar data products from multiple instruments.
- A set of naming conventions, units, metadata definitions (via HDF/HDF-EOS attributes) and data organization conventions.

Goals of Standardization

- Allow easier sharing and use of data
  - Data products which are the same between instruments have common structure and definitions
  - Identify and include attributes which aid in developing user software
  - Allow use of a single HDF/HDF-EOS library for all products
  - Provide consistent and accessible definition of data formats applicable to multiple instruments and products
- Reduce development effort and support reuse by enabling application of software to multiple instruments and products

Key Elements for developing a standard

- Identify a community of data producers and users who could benefit from standardization
- Get buy-in from every member of community
- Identify common elements which could benefit from standardization – these are the only fields you need to standardize
- Develop a consistent structure for describing data formats that can be applied even to unstandardized elements

Elements to standardize

- Dimensions and ordering of dimensions for each field
- Names of fields (includes capitalization and spacing)
- Data types for each field
- Attributes for each field, their types and definitions
  - Title, Units, MissingValue, UniqueFieldDefinition
- Units for each field (eliminates need to do conversions when using data from multiple instruments)

Results of Aura Standards

We have simplified the use and development process for both data producers and data users

- Provides defined approaches for software. Once the initial definition effort was complete developers could concentrate on science issues and not worry about how to format the data
- Assures that features that support subsequent software development are included
- Users can work with multiple parameter data sets with consistent geolocation, altitude and parameter definitions thus simplifying or eliminating much of the resampling and data conversion steps often required.

Aura Guidelines Approval Process

- Extensive use of email
  - Named authors – required to respond
  - Silent authors – could respond if desired
  - Document passed to members for extensive editing
- Controversial items were brought to individual teams for discussion by Guidelines group member(s)
- Telecons/DSWG breakout meetings held for items which required discussion
- Major releases of document voted upon by named authors

Lessons Learned

- Communicate early, before individual team's decisions on data files have been made
- Exchange data sets early on to assure common understanding of the standards
- Include software engineers and scientists in discussion group
- Be willing to compromise – will never match your "perfect data set" – remember benefits
- Group leader must be firm, no issues left unresolved
- Requires specific commitment of effort by every team and a dedicated coordinator to develop, publish and maintain documents
- Document needs to be detailed – every item not spelled out is subject to interpretation and alternate implementation

Aura Document Location

http://www.eos.ucar.edu/hirdls/HDFEOS_Aura_File_Format_Guidelines.pdf

Current Authors:

Cheryl Craig – NCAR (HIRDLS)
David Cuddy – JPL (MLS)
Scott Lewicki – JPL (TES)
Pepijn Veefkind – KNMI (OMI)
Peter Leonard – SSAI (OMI)
Al Fleig – PITA (OMI)
Paul Wagner – JPL (MLS)